The role of respiratory measures to assess mental load in pilot selection

Mariel Grassmann[a], Elke Vlemincx[b], Andreas von Leupoldt[b] and Omer Van den Bergh[b]

[a] Department of Aviation and Space Psychology, German Aerospace Center (DLR), Hamburg, Germany.
[b] Research Group on Health Psychology, University of Leuven, Leuven, Belgium.

Corresponding author:
Omer Van den Bergh
Tiensestraat 102 Box 3726
B-3000 Leuven
Tel. +32 16 32 60 58
Omer.VandenBergh@ppw.kuleuven.be

Abstract:

Whereas cardiovascular measures have a long tradition of being used to determine operator load, responsiveness of the respiratory system to mental load has rarely been investigated. In the present study we assessed basic and variability measures of respiration rate, partial pressure of end-tidal carbon dioxide (petCO_2) as well as performance measures in 63 male pilot candidates during completion of a complex cognitive task and subsequent recovery. Mental load was associated with an increase in respiration rate and a decrease in respiratory variability. A significant decrease was also found for petCO_2. Respiration rate and respiratory variability showed partial and complete effects of recovery, respectively, whereas petCO_2 did not return to baseline level. Overall, a good performance was related to a stronger reactivity in respiration rate. Our findings suggest that respiratory parameters would be a useful supplement to common measures for the assessment of mental load in pilot selection.

Keywords: mental load, multiple task performance, respiration rate, respiratory variability, end-tidal CO_2

Practitioner summary:

Respiratory measures are a promising yet poorly investigated approach to monitor operator load. For the purpose of pilot selection, we assessed respiration in response to multi-tasking in 63 candidates. Task-related changes as well as covariation with performance strongly support the consideration of respiratory parameters when evaluating reactivity to mental load.

Introduction

The quantification of operator workload is an important objective in working environments that are characterized by high task demands and responsibility. Especially in the fields of aviation, space flight, military and surgery, operational personnel are regularly faced with decision-making under time pressure and stress (e.g. Morris and Leung 2006). The aerospace sector has a long tradition of profound aptitude testing to ensure the fit between future operator and environment (Goeters, Maschke, and Eißfeldt 2004; Harris 2011). Since technical developments, however, involve constantly changing task demands impacting cognitive requirements, international airlines still seek to improve the assessment of mental load for the selection of pilots.

The construct of mental load has been defined as the ratio between task demands and operator capacity (e.g. Gopher and Donchin 1986; Kantowitz 1987; O'Donnell and Eggemeier 1986). Since human processing resources are limited, the level of mental load is influenced by difficulty and complexity of the task as well as by the individual's skills, abilities and motivation. Methods to measure mental load are generally classified on the basis of performance, self-report or psychophysiological assessment which are supposed to reflect different aspects of mental load (see Young et al. 2014). While performance-based and self-report measures are generally easier and cheaper to use, psychophysiological parameters are more cumbersome but also have some advantages over the former measures. First, they contain information about the 'physiological costs' that are related to the effort being invested to accomplish the task. In contrast to performance measures, this also allows analysts to draw conclusions about the operator's remaining capacity. Second, physiological parameters can be assessed and monitored continuously while subjective ratings are usually retrospective. Third, physiological measures are not distorted by memory lapse or observer biases. Especially in personnel selection, where the tendency to present oneself in a favourable light poses a diagnostic problem, the usefulness of subjective ratings is limited.

Whereas cardiac measures have been studied extensively as a workload index in the field of aviation (e.g. Jorna 1993; Roscoe 1993; Svensson et al. 1997; Wilson 2002) little research has been devoted to respiratory measures. The main reasons are probably that such measurements are more complicated because of the artefacts that may be caused by speech and movement as well as the fact that measurement devices assessing tidal volume are rather obtrusive. Another reason might be that respiration is in part under conscious voluntary control, while psychophysiological measures are often used because they are considered ‘objective’ measures, reflecting autonomic regulation mechanisms that are inaccessible to volitional control. Grossman and Wientjes (2001, 43) suppose, however, that ‘respiratory adjustments to highly specific behavioural demands have evolved as functional integrative adaptions to best fit and coordinate metabolic activity, cognitive performance, emotional self-regulation and perhaps even communicative signalling to conspecifics’. Awareness and conscious responsiveness of respiration are therefore crucial arguments to study the effects that psychological processes might have on the respiratory system.

The regulation of respiration is based on an interaction of the autonomic nervous system and respiratory centres in the medulla oblongata and pons on the one hand and the limbic system, cerebellum and higher cortical areas, which can affect central breathing reflexes, on the other hand. Hence, the respiratory pattern is contingent upon metabolic and homeostatic influences but also responds to cognitive and emotional processes. In contrast to the investigation of emotional influences on basic respiratory measures such as rate and volume (e.g. Bloch 1991; Boiten 1998; Boiten, Fijda, and Wientjes 1994; Gomez, Stahe, and Danuser 2004) and on variability measures of respiration (e.g. Boiten 1998; Van Diest et al. 2006; Vlemincx et al. 2013), there is only little research on the effects of mental load on human respiratory behaviour.

Respiration rate, or total breath duration, is the measure being used most frequently in respiratory reactivity research and regarded as one of the most sensitive measures of mental load (Backs, Ryan, and Wilson 1994; Brookings, Wilson and Swain 1996; Pattyn et al. 2010). Studies investigating task-related changes in respiration rate consistently show that participants breathe faster when performing a mentally demanding task both in the laboratory (Allen and Crowell 1989; Boiten 1998; Brookings et al. 1996; Mehler et al. 2009; Pattyn et al. 2010; Van Diest et al. 1999; Veltman and Gaillard 1998; Vlemincx et al. 2011; Vlemincx et al. 2012; Wientjes, Grossman, and Gaillard 1998) and in real-life settings (Harding 1987; Karavidas et al. 2010; Pattyn et al. 2010; Veltman 2002; Wilson 1993).

The analysis of respiratory variability is a promising, yet rarely applied method to investigate the psychophysiology of mental load (Vlemincx et al. 2011; Vlemincx, Van Diest, and Van den Bergh 2012; Wientjes 1992). Respiratory variability can be described as the breath-to-breath variation of spontaneous breathing and potentially reveals information about the appropriateness of respiratory control mechanisms (Tobin et al. 1988). Instead of analysing only total variability of respiratory signals, Vlemincx et al. (2010) suggested to distinguish between different components of variability, a structured or correlated component and a random component. While external influences such as cognitive and emotional demands may cause an increase in random variability, internal homeostatic regulation would cause an increase in correlated variability in order to stabilize the system. The effects of mental load on these different variability components have been investigated by means of a mental arithmetic task (Vlemincx et al. 2011). It was shown that mental load was accompanied by an increase in total variability but a decrease in correlated variability of respiration rate, suggesting an increase in the random fraction. However, in a low demanding sustained attention task, both total variability and correlated variability in respiration rate were reduced (Kagan and Rosman 1964; Vlemincx et al. 2011; Vlemincx et al. 2012). This suggests that total and correlated variability are highly sensitive to mental load, possibly differentiating between distinct cognitive processes that occur during the performance of different types of tasks.

Another respiratory measure that provides information about the fit between ventilation and metabolic demands is end-tidal carbon dioxide ($etCO_2$). The partial pressure of $etCO_2$ ($petCO_2$) is highly correlated with alveolar pCO_2 (Wientjes 1992) which, in turn, is regarded as a valid estimate of arterial pCO_2 (Phan et al. 1987; Takano et al. 2003). Given that pCO_2 decreases if ventilation exceeds metabolic demands (hyperventilation) and that respiratory behaviour is sensitive to psychological influences, it seems reasonable to investigate $petCO_2$ as a possible measure indicating over-activation in response to mental load. Wientjes et al. (1998) report a significant but small decrease in $petCO_2$ during the performance of a memory task, which is in line with other studies that investigated end-tidal and transcutaneous pCO_2 levels in response to highly demanding aviation settings (Harding 1987; Wientjes, Gaillard, and ter Maat 1996). This measure is particularly important because it has been shown that inappropriate changes in respiration rate and depth inducing hypocapnia may reduce oxygen supply to the brain and that brain oxygenation appears to be associated with impaired cognitive performance (Gibson 1978; Matthews et al. 2010; McCarthy et al. 1995; Van Diest et al. 2000). While these relationships are well established for substantial hypocapnic overbreathing, here defined as $petCO_2 < 32$ mmHg, there is little research on the covariation of task performance with less dramatically reduced $petCO_2$ levels.

Whereas inappropriate physiological arousal as indicated, for instance, by hyperventilation might impair cognitive performance, appropriate physiological arousal in response to mental load is assumed to be beneficial (Andreassi 1966; Dienstbier 1989; Lacey 1967; Jennings and Wood 1977; Obrist et al. 1974). As described by Dienstbier (1989) a low base rate of arousal, a strong sympathetic response to challenge involving release of adrenaline and noradrenaline and a quick decline of arousal after a challenge are not only associated with superior mental and physical health but also with better cognitive performance. A quick and efficient adaption to onset and offset of a demanding task indicates psychological and physiological flexibility, which generally reflects a healthy state of the individual (Kashdan and Rottenberg 2010). Flexibility is hence considered a prerequisite for the organism to properly regulate physiological arousal and to ensure optimum performance. For cardiac variability, this rationale has been supported by empirical evidence (Hansen, Johnsen, and Thayer 2003; Thayer et al. 2009).

The purpose of this study was to evaluate the sensitivity of respiratory parameters to mental load and to investigate whether task-related respiratory changes are associated with cognitive performance. In order to test this, we assessed respiration rate, variability of respiration rate and petCO_2 during a baseline period, the performance of a highly demanding cognitive multiple task, reflecting the multi-tasking demands of airline pilots, and a recovery period. We expected that respiration rate would increase from baseline to performance on the task and decrease during recovery. The total variability of respiration rate was expected to be higher, whereas the correlated fraction was expected to be suppressed while performing the multiple task. For petCO_2 we hypothesized a decrease from baseline to task and an increase during recovery. It was expected that individuals with an appropriate respiratory behaviour would perform better than individuals whose breathing pattern is not in line with metabolic demands.

Methods

Participants

The experiment was conducted at the German Aerospace Center where pilot selection is carried out on behalf of civilian airlines. 63 male volunteers were recruited from applicants who had been accepted for the selection procedure. The period of data acquisition corresponded with the ongoing selection period. In order to gain adequate power, a priori power analyses suggested a minimum of 56 participants. We continued data collection until that minimal number was surpassed. Candidates had to be fluent in German and hold a secondary-school diploma. Applicants suffering from acute illness, cardiovascular and respiratory disease or any major psychiatric disorder were not allowed to participate. They further had to be non-smokers and refrain from stimulants such as caffeine prior to their experimental testing. The data sets from two participants had to be excluded because of movement artefacts. The resulting sample consisted of 61 participants with an age range of 18 to 43 years (M = 21.8, SD = 4.2). All participants gave their written informed consent prior to the experiment and obtained 25 € for participation.

Procedure and experimental material

Experimental sessions were run around 5 pm, subsequent to the regular pilot selection tests. Participants were seated in front of a computer screen, equipped with physiological recording devices and asked to complete a questionnaire covering their age, dominant hand, regular physical activity and body mass index (BMI). After 10 to 15 minutes of signal stabilization, subjects started to go through the entire experimental protocol being guided by written instructions on a touch screen. All participants had to use their right hand to interact with the display.

The experimental protocol consisted of a six-minute resting baseline, a six-minute ‘vanilla baseline’, a six-minute multiple task and a six-minute recovery period, all of which were separated by a break of three minutes. During the resting baseline participants had to fix their eyes on a cross. A test instructor verified that participants kept their eyes focused on the monitor. The vanilla baseline was presented in addition to the resting baseline in order to reduce possible effects of anticipatory arousal with a minimally demanding vigilance task (Jennings et al. 1992). Since the presented vigilance task, however, was not accompanied by lower respiratory levels than the resting baseline, only resting baseline data were used for the statistical analysis of baseline-to-task changes. To induce mental load in a way that is similar to the cockpit workload of a civilian aircraft, a highly demanding multiple task was chosen which consisted of three single tasks measuring perceptual speed, spatial orientation and working memory capacity which had to be carried out in parallel. The perceptual speed task required the simultaneous scanning of four instruments. The pointer of each instrument indicated one of eight directions (Figure 1a). The task was to detect the number of instruments showing the same value. The spatial orientation task measured mental rotation and spatial processing ability. A pictogram of an aircraft was shown that could be rotated around its centre, pointing in one of 12 directions (comparable to a clock face). In addition, a spot was shown in one of the 12 positions on an imaginary circle around the aircraft (Figure 1b). Participants had to indicate the spot's position (clockwise 1–12) relative to the centre and flight direction of the aircraft. The working memory task required memorizing pairs of colours (grey, blue, brown, green) and two-digit numbers that were presented acoustically (e.g. 'green two eight'). A two-digit number was displayed on the screen (Figure 1c), continuously increasing from 13 to 99. The number changed every four seconds. As soon as one of the acoustically given numbers appeared on the screen, subjects had to click on the corresponding colour button. In the given example, the green button would have to be activated as soon as '28' appeared on the screen. The pairs of colour and number were generated in a controlled randomized way to assure a constant level of task difficulty. A new pair of colours and numbers was given every four seconds; the delay between acoustic and visual stimulus varied systematically between 12 and 24 seconds. The maximum number of pairs that had to be remembered simultaneously was four. Prior to the multiple task (Figure 2), participants were instructed to allocate their attention evenly to the three tasks and to work as fast and accurately as possible. During the recovery period, participants listened to relaxing music and watched an aquatic movie.

– insert Figures 1a/b/c and 2 near here –

To avoid artefacts caused by speech and movement, participants were instructed not to talk during data acquisition and to avoid any movement apart from using their right hand to operate on the touch screen. Actual movement activity was controlled by an accelerometer which was fixed to the thorax and visually monitored by the experimenter.

Performance measures

Two types of performance measures were applied: first, task performance on the experimental multiple task and second, outcome of the regular cognitive aptitude exam that was administered in the pilot selection protocol by the German Aerospace Center. Given that this test protocol covers basic mental abilities such as working memory, spatial orientation, psychomotor coordination and multi-tasking, the outcome measure is regarded as being comparable to the experimental task measure.

The performance score of the multiple task was obtained by multiplying the total number of correct responses to the three single tasks after z-standardization (i.e. sum scores were standardized across subjects within each single task). Performance data from three participants were missing due to technical problems. The dichotomous outcome measure from the cognitive aptitude exam (pass/fail) was obtained by applying the decision rules which are used for pilot selection: raw scores from the different test domains (e.g. working memory, spatial orientation, psychomotor coordination and multi-tasking abilities) were each normalized by means of a stanine transformation. Stanine scores range from 1 (low performance) to 9 (high performance) with a mean of 5 and a standard deviation of 2. If the resulting stanine scores were greater than or equal to 4 for each ability domain being tested, the cognitive aptitude exam was regarded as ‘passed’.

Physiological measures

Respiration rate (RR) and pCO_2 were measured using a mainstream capnograph (Nihon Kohden Europe GmbH, Rosbach, v.d H.) which analyses the expired air with a lightweight infrared sensor that is placed unobtrusively between nostrils and upper lip. Participants were instructed not to speak and to breathe only through their nose during the experimental periods. Respiratory data were sampled continuously at 20 Hz as well as breath-by-breath. RR was exported directly from the breath-by-breath records. The ANSLAB software (Wilhelm and Peyk 2005) was used to derive end-tidal plateau values ($petCO_2$) from the continuous pCO_2 records for each individual and period. Outliers (± 2SD) detected within one data record were corrected by linear interpolation after visual inspection.[1] Variability of RR was quantified by two types of measurement, the coefficient of variation (CV) and autocorrelation (AR). CV indicates total variability of RR within one record, whereas AR indicates the structured, correlated fraction of variability (Tobin et al. 1995). AR coefficients reported here indicate the correlation between one breath and the following one.

Prior to data analysis, the first and last 30 seconds of the six-minute data records of each experimental period were cut in order to avoid artefacts that sometimes occurred at the beginning or end of a period and to obtain stationary data which is required for the computation of AR coefficients.[2] Mean scores and variability measures were hence computed on the basis of five-minute records. Reactivity scores were obtained by calculating the difference between the mean scores of the baseline and the task period.

Statistical analysis

[1] Altogether, 5% of the 18 971 plateau values were interpolated. We confirmed our reported findings by reanalysing the $petCO_2$ data including outliers.

[2] For validation purposes, all analyses were rerun without cutting the first 30 seconds. These additional findings replicated our initial results showing that the truncated periods did not leave out crucial information.

All analyses were performed using SPSS 21.0 for Windows (SPSS Inc., Chicago, IL). To evaluate psychophysiological changes in response to mental load, respiratory variables were subjected to a repeated-measures multivariate analysis of variance (MANOVA) with period (baseline/vanilla baseline/multiple task/recovery) as a within-subject variable. Greenhouse-Geisser corrections were applied in the following ANOVAs if the assumption of sphericity was not met. If multivariate and univariate statistics were significant experimental periods were compared post-hoc using Bonferroni-corrected *t*-tests. Bivariate associations between respiratory measures and multiple task performance were assessed via Pearson's correlation. To analyse relationships with the outcome of the cognitive aptitude exam, respiratory measures were each subjected to a two-way repeated-measures ANOVA with period (baseline/vanilla baseline/multiple task/recovery) as a within- and outcome (pass/fail) as a between-subject variable. Greenhouse-Geisser corrections were applied if necessary.

We defined the family-wise alpha level as 0.05 and corrected for multiple comparisons using the Holm-Bonferroni method (Aickin and Gensler 1996; Holm 1979). Control variables (age, regular physical activity, BMI) were not included in the final analyses because they did not covary with the dependent variables. Post-hoc power analyses using GPower 3.1 confirmed that our statistical analyses were sufficiently powered (Faul, Erdfelder, Lang, and Buchner 2007).

Results

Respiratory measures and mental load

Using Pillai's trace, the MANOVA revealed a significant effect of period on RR, CV, AR and $petCO_2$ ($V = 0.82$, $F(12, 49) = 18.92$, $p < .001$). Univariate statistics are reported in Table 1. RR increased from baseline to task ($p < .001$) and decreased after the task ($p < .001$) but without going back to baseline level ($p < .001$). For CV and AR, we found a significant decrease from baseline to task and a significant increase from task to recovery period. Changes in CV were significant at a level of $p < .001$ from baseline to task and at $p < .01$ from task to recovery. Changes in AR were significant from baseline to task ($p < .05$) and from task to recovery ($p < .001$), reaching a higher level than during baseline ($p < .05$). $PetCO_2$ decreased significantly from baseline to task ($p < .05$) and did not recover after the task when comparing to baseline level (n.s.). Hypocapnic hyperventilation during task performance, here defined as $petCO_2 < 32$ mmHg, was found in three participants. In 25% of the sample we obtained $petCO_2$ levels of 34.43 mmHg or lower.

– insert Table 1 near here –

Respiratory measures and performance

Correlation analysis showed that a better performance on the experimental task was associated with lower RR during baseline ($r = -.33$, $p < .05$) and stronger reactivity in RR ($r = .34$, $p < .01$). A better task performance was also associated with higher AR during baseline ($r = .37$, $p < .01$). After excluding two outliers in the multiple task performance score (± 3 SD), resulting correlations were non-significant for RR during baseline and weaker for reactivity in RR ($r = .28$, $p < .05$) as well as for AR at baseline ($r = .30$, $p < .05$). Analyses of the relationship between respiratory measures and the outcome of the cognitive aptitude exam showed a significant interaction between experimental period and outcome of the exam for respiration rate ($F(2.19, 129.06) = 6.37$, $p < .01$, $\eta_p^2 = .10$). As depicted in Figure 3, applicants who passed the regular aptitude tests ($n = 20$) showed a higher reactivity from baseline to task and also a faster recovery after the task than those applicants who failed ($n = 41$).[i]

– insert Figure 3 near here –

Discussion

In line with our hypothesis and prior studies, respiration rate strongly increased during the performance of a difficult multiple task and decreased during the subsequent recovery period. However, respiration rate did not fully reach baseline level within the recovery time. The high sensitivity of respiration rate, as indicated by a strong effect size ($\eta_p^2 = .61$), replicates the findings reported in earlier work but using a multiple task that is comparable to the work demands of airline pilots. Our data hence further support the usefulness of respiration rate for the measurement of operator load in aviation. That respiration rate was not completely restored raises the question whether the given time window was too short in the present study. For recovery from a mental arithmetic, however, it has been shown that five to six minutes seem to be sufficient for a full return to baseline breathing frequency (Vlemincx et al. 2011). In contrast to the consistent findings on the reactivity of respiration rate to mental load, the mechanisms of recovery are poorly understood and require further investigation.

Both total and correlated variability of respiration rate were suppressed by the induction of mental load which indicates that respiratory flexibility was reduced during the task period. While a lower correlated variability is in line with existing research (Vlemincx et al. 2011; Vlemincx et al. 2012), the finding that total variability was also reduced contradicts the results reported by Vlemincx et al. (2011) for a mental arithmetic. As outlined above, a reduction in total variability has previously only been reported for sustained attention (Kagan and Rosman 1964; Vlemincx et al. 2011; Vlemincx et al. 2012). The mental task that was applied in the present experiment, however, can be characterized as attentionally highly demanding as it required allocating the necessary processing resources to the different task components. Multiple tasks, which are more appropriate when studying psychophysiological correlates of operator load in working situations that are characterized by multi-tasking, can therefore not be compared to single task tests that have been used in the mentioned studies on respiratory variability. After having finished the task, total variability returned to baseline level and correlated variability was even higher than during baseline, revealing that respiratory flexibility was fully restored within the given recovery period. The finding that the correlated fraction was higher during recovery than during baseline implies that the anticipation of an unknown multiple task might have caused an increase in random variability already during baseline. Taken together, multi-tasking – which is an important requirement in the field of aviation – appears to affect respiratory variability in a different way than single tasks.

PetCO_2, which can be regarded as an indicator of breathing in accordance with metabolic needs, decreased significantly during task performance, but only to a moderate extent. A modest but continuous hyperventilation was also reported for mental performance in the laboratory (Wientjes et al. 1998) as well as during difficult flight manoeuvers (Harding 1987) and stressful air traffic control tasks (Wientjes et al. 1996) in the field. The fact that petCO_2 reached clinically relevant levels in a small subset of participants suggests that this measure might be particularly important in order to detect individuals with a tendency to hypocapnic overbreathing, hence overreacting in highly demanding situations. Interestingly, we observed a prolonged overbreathing in this sample such that petCO_2 did not recover after task completion while respiration rate did show clear effects of recovery. This implies that tidal volume, which was not assessed in the present study, remained at a high level. In a study investigating the recovery from voluntary hyperventilation, Wilhelm, Gerlach and Roth (2001) report detailed findings on the course of petCO_2 recovery across a period of 10 minutes showing that the physiological readjustment of petCO_2 still continues after 6 minutes. In addition, tidal volume returned to baseline level within that recovery period why a time span of 10 minutes appears to be more appropriate than shorter periods to assess complete respiratory recovery.

Analysing whether respiratory behaviour is related to operator performance was a further aim of this study. The fact that three participants of the present sample responded with hypocapnic hyperventilation demonstrates that a laboratory cognitive task possibly can lead to CO_2 levels which might affect cognitive performance. An investigation of $petCO_2$ during simulated flight manoeuvres by Karavidas et al. (2010) likewise suggests that task-induced hyperventilation might be associated with performance decrements. It can be assumed that mentally demanding situations with more impact (e.g. real-life situations) would probably lead to stronger effects which then, in turn, could impair brain oxygenation. In the present study we analysed operator performance using the total performance score on the multiple task as well as the outcome of the candidates' cognitive aptitude exam that is applied in pilot selection. The analysis of respiration rate and multiple task performance revealed that lower baseline values and a strong reactivity were associated with better performance scores. Even though the task-related decrease in $petCO_2$ reported above suggests that several participants showed a breathing response exceeding metabolic demands, the present finding argues that a strong increase in respiration rate was appropriate to perform well. An explorative analysis of baseline measures revealed that correlated variability of respiration rate at rest was also associated with a better performance on the multiple task which might imply that structured respiratory variability is beneficial for cognitive processing as it is required for multiple task performance. It has to be noted, however, that the correlation coefficients were moderate and have to be interpreted with caution. When using the outcome of the cognitive aptitude exam as a performance measure, we found that candidates who passed the exam had a lower respiration rate at baseline, a stronger increase from baseline to task and again a lower respiration rate in the recovery period than candidates who failed. It could be argued that the actual exam performance might have influenced respiration rate at baseline because the experiment was conducted subsequent to the aptitude testing. Even though the candidates did not receive any feedback on their exam performance, we additionally analysed the relationship between outcome of the cognitive aptitude exam and performance on the experimental multiple task as well as respiratory variability during the experiment in order to investigate possible feedforward effects. Since none of the associations was significant in our sample, it can be assumed that respiratory behaviour during the experiment was not affected by the outcome of the cognitive aptitude testing. The reported findings thus provide additional support for a positive relationship between reactivity in respiration rate and cognitive performance. However, the present data do not allow drawing conclusions about the variables responsible for the covariation of increased ventilation and enhanced cognitive performance and future research is needed to determine the relative contribution of psychological and physiological processes to explain this relationship.

However, performance was not significantly related with petCO_2 what indicates that the amount of reactivity in petCO_2 being induced by the multiple task was too small to cause measurable effects on task performance. This interpretation is supported by previous studies suggesting that an impaired performance due to hyperventilation may occur only in individuals with a hypocapnic overbreathing response (Bloch-Salisbury, Lansing, and Shea 2000; Gibson 1978; Marangoni and Hurford 1990; Van Diest et al. 2000). As mentioned before, with the exception of three participants this was not the case in the present sample. It is further remarkable that there is a strong between-subject variance in petCO_2. Because of the presence of hypocapnic episodes in three participants and the high interindividual variation, it might be suspected that some individuals experienced some kind of emotional stress in addition to the cognitive demands. Even though our experimental design was free of any emotional cue, the mere performance situation and presence of a test instructor can possibly elicit arousal responses (see also Sonderegger and Sauer 2009).

In sum, our results replicate existing findings on the sensitivity of respiration rate towards mental load by using a task that reflects the characteristic work environment of airline pilots and, furthermore, extend the small body of research on respiratory variability under mental load. Since respiration was not only related to mental load but also to cognitive performance, we conclude that respiratory measures contain valuable information for the assessment of mental load. Including these measures in pilot selection would provide a more comprehensive picture of operator state and hence improve the diagnostic process. Future studies should compare the different respiratory measures with cardiac and other well-established measures regarding both informational content and feasibility for the purpose of application. In pilot selection, it should further be discussed whether ‘vulnerability to mental load’ as indicated by a tendency to hypocapnic overbreathing might be used as an additional criterion. Moreover, our findings provide support that resulting changes in the breathing pattern reflect not only metabolic demands but also psychological processes. However, the differentiation of cognitive and emotional influences on respiratory regulation during task performance remains an important objective for future studies.

---

[i] The dichotomous variable ‘outcome of the cognitive aptitude exam’ was composed of nine different test domains (see above). To validate these findings, we reanalysed our data using the stanine scores of those four test domains that are conceptually related to the experimental multiple task (i.e. working memory, spatial orientation, perception/concentration and multi-tasking abilities). In line with the reported results, we found significant positive correlations between performance stanine scores and the increase in RR from baseline to task, on the one hand (working memory: $r = .39$, $p < .01$; spatial orientation: $r = .36$, $p < .01$; perception/concentration: $r = .32$, $p < .05$; multi-tasking abilities: $r = .30$, $p < .05$), and the decrease in RR from task to recovery period, on the other hand (working memory: $r = .39$, $p < .01$; spatial orientation: $r = .38$, $p < .01$; perception/concentration: $r = .26$, $p < .05$; multi-tasking abilities: $r = .44$, $p < .001$).

References:

Aickin, M., and H. Gensler. 1996. “Adjusting for Multiple Testing When Reporting Research Results: The Bonferroni vs Holm Methods.” *American Journal of Public Health* 86: 726–728.

Allen, M. T., and M. D. Crowell. 1989. “Patterns of Autonomic Response During Laboratory Stressors.” *Psychophysiology* 26: 603–614.

Andreassi, J. L. 1966. “Some Physiological Correlates of Verbal Learning Task Difficulty.” *Psychonomic Science* 6: 69–70.

Backs, R. W., A. M. Ryan, and G. F. Wilson. 1994. “Psychophysiological Measures of Workload During Continuous Manual Performance.” *Human Factors* 36: 514–531.

Backs, R. W., and K. A. Seljos. 1994. “Metabolic and Cardiorespiratory Measures of Mental Effort: The Effects of Level of Difficulty in a Working Memory Task.” *International Journal of Psychophysiology* 16: 57–68.

Bloch, S., M. Lemeignan, and N. Aguilera-T. 1991. “Specific Respiratory Patterns Distinguish Among Human Basic Emotions.” *International Journal of Psychophysiology* 11: 14–154.

Bloch-Salisbury, E., R. Lansing, and S. A. Shea. 2000. “Acute Changes in Carbon Dioxide Levels Alter the Electroencephalogram Without Affecting Cognitive Function.” *Psychophysiology* 37: 418–426.

Boiten, F. H. 1998. “The Effects of Emotional Behaviour on Components of the Respiratory Cycle.” *Biological Psychology* 49: 29–51.

Boiten, F. H., N. H. Frijda, and C. J. E. Wientjes. 1994. “Emotions and Respiratory Patterns: Review and Critical Analysis.” *International Journal of Psychophysiology* 17: 103–128.

Brookings, J. B., G. F. Wilson, and C. R. Swain. 1996. "Psychophysiological Responses to Changes in Workload During Simulated Air Traffic Control." *Biological Psychology* 42: 361–377.

Dienstbier, R. A. 1989. "Arousal and Physiological Toughness: Implications for Mental and Physical Health." *Psychological Review* 96: 84–100.

Faul, F., E. Erdfelder, A.-G. Lang , and A. Buchner. 2007. "G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences." *Behavior Research Methods* 39: 175–191.

Gibson, T. M. 1978. "Effects of Hypocapnia on Psychomotor and Intellectual Performance." *Aviation, Space and Environmental Medicine* 49: 943–946.

Goeters, K.-M., P. Maschke, and H. Eißfeldt. 2004. "Ability Requirements in Core Aviation Professions: Job Analysis of Airline Pilots and Air Traffic Controllers." In *Aviation Psychology: Practice and Research*, edited by K.-M. Goeters, 99–119. Aldershot: Ashgate Publishing.

Gomez, P., W. A. Stahe, and B. Danuser. 2004. "Respiratory Responses During Affective Picture Viewing." *Biological Psychology* 67: 359–373.

Gopher, D., and E. Donchin. 1986. "Workload: An Examination of the Concept." In *Handbook of Perception and Human Performance*, edited by K. R. Boff, L. Kaufman. and J. P. Thomas, 41-1–41-49. Oxford: John Wiley & Sons.

Grossman, P., and C. J. E. Wientjes. 2001. "How Breathing Adjusts to Mental and Physical Demands." In *Respiration and Emotion*, edited by Y. Haruki, I. Homma, A. Umezawa, and Y. Masaoka, 43–54. Tokyo: Springer.

Hansen, A. L., B. H. Johnsen, and J. Thayer. 2003. "Vagal Influence on Working Memory and Attention." *International Journal of Psychophysiology* 48: 263–274.

Harding, R. M. 1987. *Human Respiratory Responses During High Performance Flight*. AGARDograph No. 312W, NATO Advisory Group for Aerospace Research and Development, 1–83.

Harris, D. 2011. *Human Performance on the Flight Deck*. Farnham: Ashgate Publishing.

Holm, S. 1979. “A Simple Sequentially Rejective Multiple Test Procedure.” *Scandinavian Journal of Statistics* 6: 65–70.

Jennings, J. R., T. Kamarck, C. Stewart, M. Eddy, and P. Johnson. 1992. “Alternate Cardiovascular Baseline Techniques: Vanilla or Resting Baseline.” *Psychophysiology* 29: 742–750.

Jennings, J. R., and C. C. Wood. 1977. “Cardiac Cycle Time Effects on Performance, Phasic Cardiac Responses and Their Intercorrelation in Choice Reaction Time.” *Psychophysiology* 14: 297–307.

Jorna, P. G. A. M. 1993. “Heart Rate and Workload Variations in Actual and Simulated Flight.” *Ergonomics* 36: 1043–1054.

Kagan, J., and B. L. Rosman. 1964. “Cardiac and Respiratory Correlates of Attention and an Analytic Attitude.” *Journal of Experimental Child Psychology* 1: 50–63.

Kantowitz, B. H., and P. A. Caper. 1988. “Human Workload in Aviation.” In *Human Factors in Aviation*, edited by E. Wiener, and D. Nagel, 157–187. New York: Academic Press.

Karavidas, M. K., P. M. Lehrer, S. E. Lu, E. Vaschillo, B. Vaschillo, and A. Cheng. 2010. “The Effects of Workload on Respiratory Variables in Simulated Flight: A Preliminary Study.” *Biological Psychology* 84: 157–60.

Kashdan, T. B., and J. Rottenberg. 2010. “Psychological Flexibility as a Fundamental Aspect of Health.” *Clinical Psychology Review* 30: 865–878.

Lacey, J. L. 1967. “Somatic Response Patterning and Stress: Some Revisions of Activation Theory.” In *Psychological Stress: Issues in Research*, edited by M. H. Appley, and R. Trumbull, 14–42. New York: Appleton-Century-Crofts.

Marangoni, A. H., and D. P. Hurford. 1990. “The Effect of Varying Alveolar Carbon Dioxide Levels on Free Recall.” *Brain and Cognition* 13: 77–85.

Matthews, G., J. S. Warm, L. E. Reinerman-Jones, L. K. Langheim, D. A. Washburn, and L. Tripp. 2010. “Task Engagement, Cerebral Blood Flow Velocity, and Diagnostic Monitoring for Sustained Attention.” *Applied Journal of Experimental Psychology* 16: 187–203.

McCarthy, D., R. Corban, S. Legg, and J. Faris. 1995. “Effects of Mild Hypoxia on Perceptual-motor Performance: A Signal-detection Approach.” *Ergonomics* 38: 1979–1992.

Mehler, B., B. Reimer, J. F. Coughlin, and J. A. Dusek. 2009. “Impact of Incremental Increases in Cognitive Workload on Physiological Arousal and Performance in Young Adult Drivers.” *Transportation Research Record: Journal of the Transportation Research Board* 2138: 6–12.

Morris, C. H., and Y. K. Leung. 2006. “Pilot Mental Workload: How Well Do Pilots Really Perform?” *Ergonomics* 49: 1581–1596.

O’Donnell, R. D., and F. T. Eggemeier. 1986. “Workload Assessment Methodology.” In *Handbook of Perception and Human Performance: Vol. II. Cognitive Processes and Performance*, edited by K. R. Boff, L. Kaufman, and J. P. Thomas, 42-1–42-49. New York: Wiley.

Obrist, P. A., J. L. Howard, J. E. Lawler, R. A. Galosy, K. A. Meyers, and C. J. Gaebelein. 1974. “The Cardiac-somatic Interaction.” In *Cardiovascular Psychophysiology: Current Issues in Response Mechanisms, Biofeedback, and Methodology*, edited by P. A. Obrist, A. H. Black, J. Brener, and L. V. DiCara, 136–162. Chicago: Aldine.

Pattyn, N., P.-F. Migeotte, X. Neyt, A. van den Nest, and R. Cluydts. 2010. "Comparing Real-life and Laboratory-induced Stress Reactivity on Cardio-respiratory Parameters: Differentiation of a Tonic and a Phasic Component." *Physiology & Behaviour* 101: 218–223.

Phan, C. Q., K. K. Tremper, S. E. Lee, and S. J. Barker. 1987. "Noninvasive Monitoring of Carbon Dioxide: A Comparison of the Partial Pressure of Transcutaneous and End-tidal Carbon Dioxide With the Partial Pressure of Arterial Carbon Dioxide." *Journal of Clinical Monitoring* 3: 149–154.

Roscoe, A. H. 1993. "Heart Rate as a Psychophysiological Measure for In-flight Workload Assessment." *Ergonomics* 36: 1055–1062.

Sonderegger, A., and J. Sauer. 2009. "The Influence of Laboratory Set-up in Usability Tests: Effects on User Performance, Subjective Ratings, and Physiological Measures." *Ergonomics* 52: 1350–1361.

Svensson, E., M. Angelborg-Thanderz, L. Sjöberg, and S. Olsson. 1997. "Information Complexity – Mental Workload and Performance in Combat Aircraft." *Ergonomics* 40: 362–380.

Takano, Y., O. Sakamoto, C. Kiyofuji, and K. Ito. 2003. "A Comparison of the End-tidal CO_2 Measured by Portable Capnometer and the Arterial pCO_2 in Spontaneously Breathing Patients." *Respiratory Medicine* 97: 476–481.

Thayer, J. F., A. L. Hansen, E. Saus-Rose, and B. H. Johnsen. 2009. "Heart Rate Variability, Prefrontal Neural Function, and Cognitive Performance: The Neurovisceral Integration Perspective on Self-regulation, Adaptation, and Health." *Annals of Behavioural Medicine* 37: 141–153.

Tobin, M. J., M. J. Mador, S. M. Guenther, R. F. Lodato., and M. A. Sackner. 1988. “Variability of Resting Respiratory Drive and Timing in Healthy Subjects.” *Journal of Applied Physiology* 65: 309–317.

Tobin, M. J., K. L. Yang, A. Jubran, and R. F. Lodato. 1995. “Interrelationship of Breath Components in Neighboring Breaths of Normal Eupnic Subjects.” *American Journal of Respiratory and Critical Care Medicine* 152: 1967–1976.

Van Diest, I., K. Stegen, K. P. Van de Woestijne, N. Schippers, and O. Van den Bergh. 2000. “Hyperventilation and Attention: Effects of Hypocapnia on Performance in a Stroop Task.” *Biological Psychology* 53: 133–252.

Van Diest, I., K. Stegen, K. P. Van de Woestijne, and O. Van den Bergh. 1999. “Critical Conditions for Respiratory Stress Responses: A Literature Review.” [In Dutch.] *Gedrag & Gezondheid* 27: 290–303.

Van Diest, I., J. F. Thayer, B. Vandeputte, K. P. Van de Woestijne, and O. Van den Bergh. 2006. “Anxiety and Respiratory Variability.” *Physiology & Behaviour* 89: 189–195.

Veltman, J. A. 2002. “A Comparative Study of Psychophysiological Reactions During Simulator and Real Flight.” *The International Journal of Aviation Psychology* 12: 33–48.

Veltman, J. A., and A. W. K. Gaillard. 1998. “Physiological Workload Reactions to Increasing Levels of Task Difficulty.” *Ergonomics* 41: 656–669.

Vlemincx, E., J. Taelman, S. De Peuter, I. Van Diest, and O. Van den Bergh. 2011. “Sigh Rate and Respiratory Variability During Mental Load and Sustained Attention.” *Psychophysiology* 48: 117–20.

Vlemincx, E., I. Van Diest, P. M. Lehrer, A. E. Aubert, and O. Van den Bergh. 2010. “Respiratory Variability Preceding and Following Sighs: A Resetter Hypothesis”. *Biological Psychology* 84: 82–87.

Vlemincx, E., I. Van Diest, and O. Van den Bergh. 2012. “A Sigh Following Sustained Attention and Mental Stress: Effects on Respiratory Variability.” *Physiology & Behaviour* 107: 1–6.

Vlemincx, E., D. Vigo, D. Vansteenwegen, O. Van den Bergh, and I. Van Diest. 2013. “Do Not Worry, Be Mindful: Effects of Induced Worry and Mindfulness on Respiratory Variability in a Nonanxious Population.” *International Journal of Psychophysiology* 87: 147–151.

Wientjes, C. J. E. 1992. “Respiration in Psychophysiology: Measurement Issues and Applications.” *Biological Psychology* 34: 179–203.

Wientjes, C. J. E., A. W. K. Gaillard, and R. ter Maat. 1996. “Measurement of Transcutaneous pCO_2 Among Air Traffic Controllers: Differential Responding to Workload and Stress.” *Biological Psychology* 43: 250–251.

Wilhelm, F. H., A. L. Gerlach, and W. T. Roth. 2001. “Slow Recovery from Voluntary Hyperventilation in Panic Disorder.” *Psychosomatic Medicine* 63: 638–649.

Wientjes, C. J. E., P. Grossman, and A. W. K. Gaillard. 1998. “Influence of Drive and Timing Mechanisms on Breathing Pattern and Ventilation During Mental Task Performance.” *Biological Psychology* 49: 53–70.

Wilson, G. F. 1993. “Air-to-ground Training Missions: A Psychophysiological Workload Analysis.” *Ergonomics* 36: 1071–1087.

Wilson, G. F. 2002. “An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures.” *The International Journal of Aviation Psychology* 12: 3–18.

Young, M. S., K. A. Brookhuis, C. D. Wickens, and P. A. Hancock. 2014. “State of Science: Mental Workload in Ergonomics.” *Ergonomics* doi: 10.1080/00140139.2014.956151.

Table 1. Means (*M*), standard deviations (SD) and repeated-measures ANOVA results (including effect size η_p^2) for basic and variability respiratory parameters.

		BL		V-BL		MT		RC						
	N	*M*	SD	*M*	SD	*M*	SD	*M*	SD	df	df_{error}	*F*	*p*	η_p^2
RR	61	13.24[a]	3.95	14.90[b]	3.65	18.96[c]	3.15	14.64[b]	4.00	2.05	123.02	80.91	<.001	.57
CV (RR)	61	.19[a]	.09	.17[a]	.07	.13[b]	.05	.19[a]	.13	2.49	149.33	12.71	<.001	.18
AR (RR)	61	.13[a]	.17	.14[a]	.16	.05[b]	.11	.21[a]	.16	3	180	12.48	<.001	.17
$petCO_2$	61	36.80[a]	2.98	36.59[a]	2.70	36.06[b]	2.75	35.72[c]	3.02	2.26	135.70	12.56	<.001	.17

RR: respiration rate, CV: coefficient of variation, AR: autocorrelation, $petCO_2$: partial pressure of end-tidal carbon dioxide, BL: baseline, V-BL: vanilla baseline, MT: multiple task, RC: recovery.

[a, b, c] Different letters in the same row indicate significant differences ($p < .05$).

Figure captions:

Figure 1. Exemplary items and corresponding keypads of the (a) perceptual speed task (correct response: ‘2’), (b) spatial orientation task (correct response: ‘2’) and (c) working memory task.

Figure 2. Screenshot of the multiple task.

Figure 3. Mean respiration rate during baseline, multiple task and recovery for participants who passed (*n* = 20) and participants who failed (*n* = 41) the cognitive aptitude exam for pilot selection. Error bars indicate standard deviations.